

Software Engineering Plan for the Land Information System

Submitted under Task Agreement GSFC-CT-2

Cooperative Agreement Notice (CAN) CAN-00OES-01

**Increasing Interoperability and Performance of Grand Challenge
Applications in the Earth, Space, Life, and Microgravity Sciences**

Version 2.1

October 6, 2003

Revision	Changes	Date
2.1	Milestone I	September 15, 2003
1.0	Milestone A	June 26, 2002

Table of Contents

Table of Contents	2
1. Introduction	3
1.1 Purpose	3
1.2 Background	3
1.3 Organization and Responsibilities	3
1.3.1 Development Team	3
1.3.2 Supporting Team	4
1.3.3 Interfacing Groups	5
2. LIS Problem Statement	6
3. Technical Approach	7
3.1 Assumptions and Constraints	7
3.2 Development Environment	8
3.3 Activities, tools, and products	9
3.4 Build Strategy	11
3.5 Code Re-use	14
4. Management Approach	14
4.1 Assumptions and Constraints	14
4.2 Milestones and Schedules	15
4.3 Metrics	15
4.4 Risk Management	16
5. Product Assurance	17
5.1 Configuration Management	17
5.2 Quality Assurance	18

1. Introduction

This document describes the Software Engineering Plan for the Land Information System (LIS) to be implemented under funding from NASA's ESTO Computational Technologies (formerly High Performance Computing and Communications) Project. This plan is submitted to satisfy the Task Agreement GSFC-CT-2 under Cooperative Agreement Notice CAN-00-OES-01 Increasing Interoperability and Performance of Grand Challenge Applications in the Earth, Space, Life, and Microgravity Sciences.

1.1 Purpose

The Land Information System (LIS) project is to develop a leading-edge land surface modeling and data assimilation system, to support broad land surface research and application activities, to help define Earth system modeling interoperability standards, and to lead the effective application of Internet technology and cost-effective parallel processing to high-resolution, real-time Earth system studies.

1.2 Background

The Land Information System (LIS) is to be built upon the existing Land Data Assimilation Systems (LDAS), a pilot research effort of land surface modeling and data assimilation. LDAS has demonstrated successful land modeling capability and usability. However, LDAS can only operate with moderate spatial resolutions and no parallel processing support. This makes it impractical for potential high-resolution, real-time applications. LIS will inherit LDAS' land modeling and data assimilation capability, while adding parallel processing support, a web-based user interface, mass data storage, interoperability and portability, to push the envelope of the current land modeling effort towards the goal of making a high-resolution, real-time, operation-quality land information system. The Land Information System will enhance land surface research and applications and help define land surface modeling and assimilation standards. LIS will also assist in the definition and demonstration of the Earth System Modeling Framework (ESMF).

1.3 Organization and Responsibilities

The Hydrological Sciences Branch at NASA's Goddard Space Flight Center (GSFC) manages the LIS. The LIS development personnel are organized as two teams: the development team and the supporting team. Contact information for the team is available from <http://lis.gsfc.nasa.gov/contacts/>.

1.3.1 Development Team

The development team is responsible for the overall project management, team coordination, and development, maintenance, port and support of the core LIS system. The development team mainly consists of the team members at NASA/GSFC, with the following responsibilities:

- Dr. P. Houser: scientific supervision

- Dr. C. Peters-Lidard: project management, milestone enforcement, coordination, etc.
- Ms. S. Olden: System documentation, software engineering, quality assurance.
- Dr. S. Kumar and Mr. J. Geiger: LDAS analysis and documentation, parallelization of LIS on Origin and Linux Cluster, configuration management of the software.
- Dr. Y. Tian: LIS hardware and software analysis and design, including data management. Leads a team of assemblers and system administrators.
- Mr. L. Lighty: Web interface design and implementation. Webmaster.

Drs. Houser and Peters-Lidard are serving as co-PI's on the project. They have substantial expertise in land modeling and the application of HPCC technologies to land-atmosphere modeling. More details about their background and experience are given below.

P. R. Houser (PI; NASA/GSFC) Dr. Houser's area of expertise is local to global land surface-atmospheric observation and numerical simulation, development and application of hydrologic data assimilation methods, and multi-scale soil moisture investigations. Dr. Houser's primary responsibility will be the management and coordination of the LIS team.

C. D. Peters-Lidard (Co-PI; NASA/GSFC) Dr. Peters-Lidard focuses on measurement and modeling of terrestrial water and energy balances and fluxes due to land-atmosphere interactions over a range of temporal and spatial scales. Dr. Peters-Lidard has extensive experience with massively parallel, coupled land-surface atmospheric models, and will have a primary responsibility in this area. Further, she will facilitate many aspects of clustered high performance computing, atmospheric coupling, earth system model connections, and advanced software engineering. Dr. Peters-Lidard will direct 5 full time computer scientist and software engineers working on LIS.

1.3.2 Supporting Team

The supporting team will supervise scientific issues, interoperability, land model development and data management utilities.

- Dr. P. Dirmeyer: ALMA, interface/coupling standards
- Mr. B. Doty: GrADS and GrADS-DODS Server
- Dr. K. Mitchell: NOAH Land Surface Model
- Dr. E. Wood and J. Sheffield: VIC Land Surface Model
- Dr. S. Denning: Advisor on future carbon modeling

P. Dirmeyer (co-I; COLA) Dr. Dirmeyer conducts research into the role of land surface in the climate system. This includes development and application of land-surface models, studies of the impact of the land on the predictability of climate, and the impacts of land use change on regional and global climate. Dr. Dirmeyer is chair of the GEWEX global soil wetness project and as such has developed extensive community infrastructure frameworks and tools, such as the ALMA land surface model standards. Dr. Dirmeyer will be responsible for LIS software standards, the Community Land Model framework,

will provide input to atmosphere coupling, earth system models, and software engineering. Dr. Dirmeyer will also serve as a collaboration mechanism with GEWEX global soil Wetness Project, ISLSCP, and COLA.

B. Doty (Co-I;COLA) Brian Doty has over 20 years of in-depth experience in a wide range of computing technologies. He developed the Grid Analysis and Display System (GrADS) package that is used worldwide for climate data analysis and visualization. He will be primarily responsible for the application of the GrADS-DODS Server (GDS) software to the LIS, in particular its support of remote users who wish to access and analyze high resolution land surface model output and remote sensing data sets. He will also initiate the work toward utilizing parallel data analysis techniques from within GrADS to permit rapid access and computation on the large data sets envisioned in this project.

K. Mitchell (Co-I; NCEP) Dr. Mitchell is primarily responsible for land modeling and its coupling to all regional and global weather and climate prediction models, and their supporting data assimilation systems, at the National Center for Environmental Prediction. He is the major developer of the NOAA land model, its NWP connections, and is a leader in the LDAS community framework. As part of an operational center, Dr. Mitchell has extensive experience with efficient land model coupled code implementation on very large compute platforms. Dr. Mitchell will be responsible for providing the NOAA model and guidance on its execution in the LIS, including supporting documentation, ancillary files, example control runs, and cold-start initialization strategies. Additionally, Dr. Mitchell will advise on LIS atmosphere connections and provide feedback from the North American LDAS project, including validation and assimilation approaches, as well as software utilities.

E. Wood (Co-I;Princeton) Professor Wood's specialty is hydroclimatology, with research focuses in land-atmospheric interactions, terrestrial remote sensing, hydrologic impacts from climate change, and environmental data analysis. Professor Wood will be responsible for the inclusion of the VIC model in LIS, various aspects of LIS remote sensing, and general hydrologic modeling and analysis. Professor Wood has one full time computer scientist working on the VIC model and LIS implementation of the model.

S. Denning (Co-I; Colorado State University) Professor Denning investigates the interface between terrestrial ecosystems and the atmosphere, with particular emphasis on using atmospheric observation to understand the global carbon cycle. His research involves the development and application of simulation models of these constraints on the carbon budget of the atmosphere. He emphasized the "fusion" of modeling and data analysis to understand natural processes. Professor Denning will be advise on the land surface biogeochemistry and carbon cycle remote sensing and modeling aspects of LIS.

1.3.3 Interfacing Groups

LIS will interface with a variety of groups in addition to the Global Land Data Assimilation System (GLDAS), the North American Land Data Assimilation System (NLDAS), the GrADS/DODS, and the Community Land Model community, the NOAA

Land Model community and the VIC Land Model community. As noted above, leaders of the GLDAS, NLDAS, GrADS/DODS, CLM, NOAH and VIC communities are part of the LIS team, and regularly participate in group meetings and monthly teleconferences. Other interfacing groups will include the Earth System Modeling Framework (ESMF) community, the NASA ESTO Computational Technologies Project, John Dorband's Linux cluster group at GSFC, and the ALMA standards group. Interaction with the ESMF community will be accomplished via formal participation in each other's Review Boards (Arlindo da Silva on the LIS board, and Paul Houser on the ESMF board), in addition to informal interaction at ESMF meetings, and via email and/or phone. Interaction with the CT project will be coordinated with respect to milestone review and on an as-needed basis. We have already held several meetings with John Dorband's Linux cluster group, and will continue to do so as needed. Finally, interaction with the ALMA standards group will be accomplished via the appointment of the PRISM/ALMA Land Surface Modeling Lead, Dr. Jan Polcher, to our Review Board.

2. LIS Problem Statement

The Land Information System (LIS) will have the following components: (1) A high-resolution (1km) global Land Data Assimilation System (LDAS), involving several independent community Land Surface Schemes (LSS), land surface data assimilation technologies, and integrated database operations for observation and prediction data management; and (2) A web-based user interface based on the Grid Analysis and Display System (GrADS) and the Distributed Oceanographic Data System (DODS) for accessing data mining, numerical modeling, and visualization tools. The LIS will be available as a "production" system implemented on the custom LIS Linux cluster for high-resolution global applications, and will also be tested and demonstrated on SGI Origin 3000 systems for portability and potential smaller, community applications. By incorporating and promulgating the existing Assistance for Land surface Modeling Activities (ALMA; <http://www.lmd.jussieu.fr/ALMA/>) and DODS standards for model coupling and data visualization, LIS will contribute to the definition of the land surface modeling and assimilation standards for the Earth System Modeling Framework (ESMF).

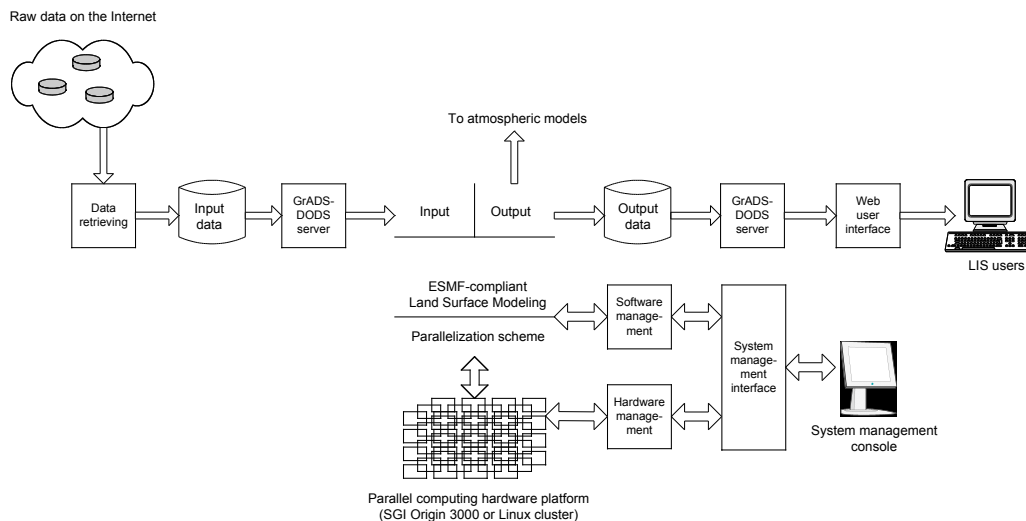


Figure 1. LIS System Architecture

LIS has to face challenges in the following six areas,

- **Modeling reality.** This is a test of our scientific understanding of the physical processes of the global land surface. Fortunately, the LIS developer team has the leading researchers in the land surface modeling community, and the scientific issues have been attacked in the existing LDAS project extensively, so we are confident that the LIS will be built on a solid scientific foundation. Moreover, the development of LIS and its high-quality data will help us to advance our scientific research on land modeling.
- **Performance.** The requirements of near real-time, high-resolution demands high performance computing. From base-lining the LDAS runs, we have obtained realistic estimate of the required computing power for LIS, and have gained insight on efficient parallelization schemes we will use for LIS. We will both build a cost-effective Linux cluster, and use NASA's existing Origin supercomputers, with parallelized LIS code, to meet the demand.
- **Data storage.** The high-resolution LIS system will produce a huge amount of data. To manage this issue, we will take advantage of our Linux cluster's distributed storage feature, to store output data in a clustered environment to ensure reliable, high-performance data management operations.
- **Usability.** LIS will be a sophisticated system. To let users access LIS, we will design intuitive web-based interfaces, and users can have different levels of access, depending on their needs. Meanwhile, we will extensively employ GrADS's graphical capability and GrADS/DODS server's data sub-setting functionality to enable users efficient access and use of the LIS data.
- **Interoperability.** LIS is required to be able to interface with other Earth system models. This can be achieved by incorporating the ALMA (Assistance for Land surface Modeling Activities) standard, and by following and contributing to the ESMF (Earth System Modeling Framework) standard.
- **Portability.** LIS has to be able to run both on NASA's Origin supercomputers, and on our custom-built Linux cluster. This is easy to achieve, since the LIS source code for scientific computing will be written in the high-level languages, Fortran and C, with MPI libraries. Since Fortran, C and MPI are all highly standard and portable, as long as we refrain from using any platform-specific language features, we will deliver LIS as a portable system.

In summary, developing LIS presents quite a few challenges, and we are confident that our skills, experience and development strategy will be able to handle them to achieve our goal, and in the meantime, contribute to the knowledge and advancement in these areas.

3. Technical Approach

3.1 Assumptions and Constraints

This plan is based on, and constrained by the following factors:

- Our scientific understanding of the land surface physics and processes. We are under the assumption that based on the current state of the art science understanding of the macroscale water, energy and momentum transfer processes,

and we will be able to realistically reproduce the Earth's land surface and its dynamics in our computer models. However, the land surface processes are highly complicated, and our success in this aspect is limited by the current scientific knowledge in this area.

- The development, validation and operation of LIS will heavily rely on the availability and quality of the forcing and observation data. These data fundamentally dictate the LIS output data.
- Sufficient and dedicated computing resources. We expect the NASA/ARC Origin supercomputing facility will provide us with a suitable development environment and ample computing power. Also the LIS development and operation depends critically on the successful installation and smooth operation of our Linux cluster within the constraint of the budget and human resources.

3.2 Development Environment

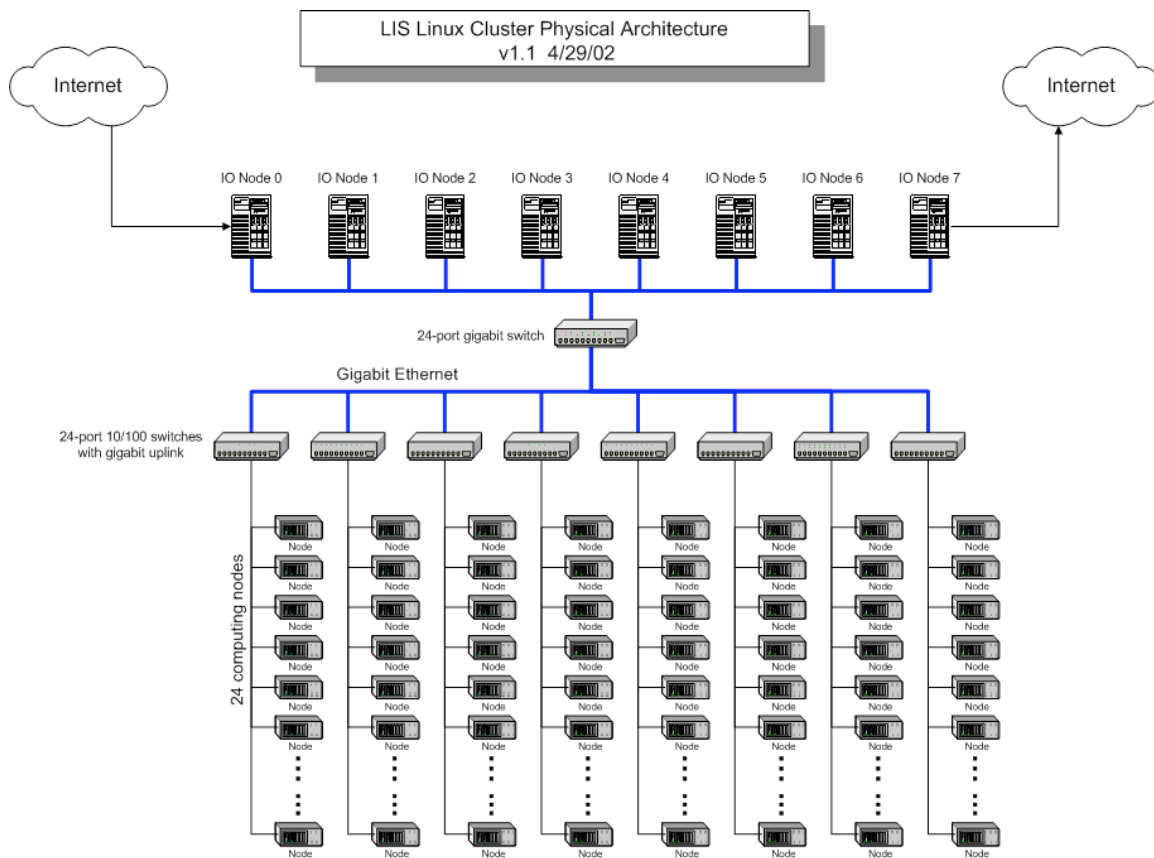


Figure 2. LIS Linux Cluster Architecture

The primary hardware platform will be a Linux cluster with approximately 200 PC nodes, running standard Linux kernel. A secondary hardware platform is SGI Origin 2000 (and eventually SGI Origin 3000) running IRIX, used as a test bed for prototyping and verification of software portability.

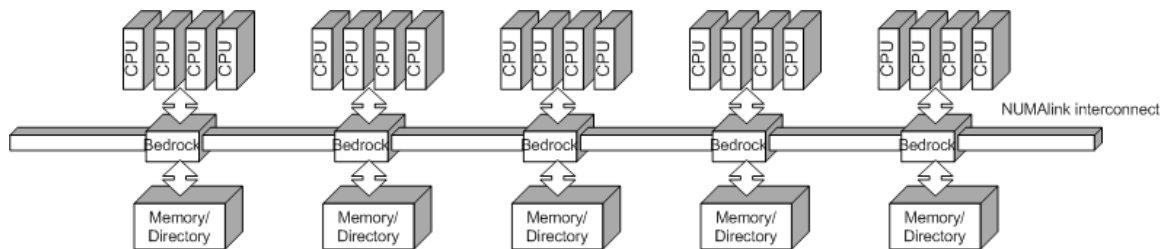


Figure 3. SGI Origin 3000 Architecture

Scientific computation of LIS on both platforms will be developed mainly in FORTRAN 90 and C. Parallel computing will be implemented with MPI. System management and job scheduling will be done with third-party open source Unix/Linux packages, supplemented with software utilities developed in-house. User interfaces will be largely web-based, developed in static HTML or interactive CGI scripts, with Cascade Style Sheets (CSS) for page layout.

3.3 Activities, tools, and products

The development activities will revolve around the standard software development cycle: requirement analysis, system design, developing, testing and maintenance:

- Requirements analysis. The LIS requirements analysis process has involved initial input from the Development Team, based on detailed performance analysis and baselining of the existing LDAS system, in addition to system-level requirements based on the LIS system concept, including the user interface. The performance and resource usage information gathered will enable us to predict the requirements for LIS. The product of the requirement analysis will be LDAS performance data and LDAS source code documentation, as well as LIS requirement projections resulting from this analysis. In addition, the implication of LIS' real-time, high-resolution requirements will also be analyzed and documented. Community input has been and will continue to be solicited during teleconferences with the Supporting Team and Review Board. The initial analysis will be submitted via the web before project milestone (E). System design. We will employ a systematic, top-down approach for the design of the system. We will start with the top level system physical and logical architecture, working our way down the various system components and sub-components. Industry standard design tools, such as Microsoft Visio, will be used. The system design will produce detailed diagrams and specifications of the system, including hardware layout and connection, software structure and flow chart, network design, and data management infrastructure. Prototyping will be used especially in the area of parallelization. The parallelization scheme will be prototyped on both SGI Origin and the Linux Cluster. The performance of parallel versions of the code will be compared with expected performance benchmarks. The scalability of the performance scheme will be tested by applying it to different resolutions of global domain. The performance studies

will be conducted for different land surface models as well. Prototyping will also be used during the design of the interoperability scheme.

The result of the design will be detailed design documentation in non-proprietary format, and will be submitted to meet project milestone (E).

- **Development.** As discussed further below, the LIS is based largely on public-domain land surface models, database and visualization tools and the existing GLDAS. These codes have been developed at their respective institutions, each with their own standard coding practice, and will remain under these practices for the duration of the LIS project. We will follow GSFC standard coding practices and styles with respect to the GLDAS and LIS user interface development, and all code will be open-source and freely available. The source code will be extensively documented both in-line and stand-alone. The source code will be accompanied by detailed documentation, including industry-standard diagrams such as calling-tree diagrams and data flow diagrams. All the development and code improvement will be delivered at the corresponding LIS milestones.
- **Testing.** We will carry out the following testing activities at every proper stage of the development cycle:
 - a. **Component testing.** Each component will be tested to ensure validity and optimal performance.
 - b. **Integration testing.** We will make sure all the pieces can work together, produce correct results and achieve maximum performance.
 - c. **Validation testing.** The system output will be tested against known output from LDAS and compare with observational data.
 - d. **Job stream testing.** At every stage of the development, we will make sure the system will work smoothly in a production environment, from data input, computation, data analysis, data output all the way to user interaction.
 - e. **Interface testing.** Tests will be done to make sure LIS will interface with the data suppliers, the users and land surface models consistently.
 - f. **Security testing.** The system will be tested against real-world situations where incorrect input, user errors, malicious users, software bugs, to correct any potential security hazards.
 - g. **Stability and recovery testing.** The system will be tested to verify great fault tolerance and graceful recovery.
 - h. **Performance testing.** Performance will be measured on both platforms to see if the performance requirements are met, and provide feedback for bottleneck identification and fine tuning.
 - i. **System Level testing.** System level testing will be performed after the system is completely integrated. Tests will be designed to verify requirements.
 - j. **Regression testing.** Regression testing is performed to ensure that capabilities previously provided remain intact. Regression testing will be performed after each build is delivered.

The initial version of the test plan will be delivered via web at LIS milestone (H). A detailed testing report will be produced at every milestone to guide further development and will be posted to the LIS website.

- **Maintenance.** Long-term maintenance of the LIS system is expected to be performed by the GLDAS team, and will be facilitated by the delivery of a User's Guide (Milestone I) and Maintenance Manual (Milestone K). Depending on the availability of funds for future LIS development, including possible coupling to other Earth System Models, the LIS team may continue to maintain the LIS code. Either the GLDAS or the LIS staff will maintain the bug tracking and reporting tools developed for LIS.

3.4 Build Strategy

The LIS build strategy follows the milestones set forth for the project. Table 1 below illustrates the five code delivery milestones and associated builds, including the baselining of the original code. Each build is described in more detail below.

Table 1. LIS milestones and build strategy.

Name		Start	Finish
LIS Baselining	143.d	1/14/02 8:00	7/31/02 17:00
Clean Code	60.d	1/14/02 8:00	4/5/02 17:00
Run Code	60.d	4/8/02 8:00	6/28/02 17:00
LIS Baseline Code Build	.d	7/31/02 17:00	7/31/02 17:00
Build 1: First Code Improvement	173.d	8/1/02 8:00	3/31/03 17:00
Implement ESMF Time Management	60.d	8/1/02 8:00	10/23/02 17:00
Build 0.1: ESMF Time MGT in LDAS	.d	10/23/02 17:00	10/23/02 17:00
Parallelize LDAS	90.d	10/24/02 8:00	2/26/03 17:00
Build 0.2: Parallel LDAS	.d	2/26/03 17:00	2/26/03 17:00
Prepare 5-km Data & Test	60.d	1/1/03 8:00	3/25/03 17:00
Build 0.3: 5-km Data Test	.d	3/25/03 17:00	3/25/03 17:00
Prepare Final Documentation	75.d	8/1/02 8:00	11/13/02 17:00
Prepare Final Build	4.d	3/26/03 8:00	3/31/03 17:00
Build 0.4 Final Build & Test	150.d	8/1/02 8:00	2/26/03 17:00
Build 1: First Code Improvement	.d	3/31/03 17:00	3/31/03 17:00
F) First code improvement completed	.d	3/31/03 17:00	3/31/03 17:00
Build 2: Interoperability Prototype	88.d	4/1/03 8:00	7/31/03 17:00
Implement ESMF Utility/Clean up/Port to Cluster	20.d	4/1/03 8:00	4/28/03 17:00
Build 1.1: ESMF Utility/Clean up/Port to Cluster	.d	4/28/03 17:00	4/28/03 17:00
Add GLDAS Functionality	20.d	4/29/03 8:00	5/26/03 17:00
Build 1.2: GLDAS/NCEP Functionality	.d	5/26/03 17:00	5/26/03 17:00
Incorporate VIC model	30.d	4/29/03 8:00	6/9/03 17:00
Prepare VIC 5-km datasets	15.d	4/1/03 8:00	4/21/03 17:00

Run and test VIC 5-km datasets	10.d	6/10/03 8:00	6/23/03 17:00
Build 1.3: VIC Model	.d	6/23/03 17:00	6/23/03 17:00
Implement GDS-based I/O, CLM/Noah to Cluster	60.d	4/1/03 8:00	6/23/03 17:00
Build 1.4: GDS-based I/O	.d	6/23/03 17:00	6/23/03 17:00
Implement ALMA-based wrapper	5.d	6/24/03 8:00	6/30/03 17:00
Prepare Final Documentation	75.d	4/1/03 8:00	7/14/03 17:00
Prepare Final Build	5.d	7/15/03 8:00	7/21/03 17:00
Build 2.0 Final Build & Test	.d	7/21/03 17:00	7/21/03 17:00
Build 2: Interoperability Prototype	.d	7/31/03 17:00	7/31/03 17:00
I) Interoperability Prototype	.d	7/31/03 17:00	7/31/03 17:00
Build 3: Second Code Improvement	140.d	8/1/03 8:00	2/27/04 17:00
Prepare NOAH 1-km datasets	10.d	8/1/03 8:00	8/14/03 17:00
Prepare CLM 1-km datasets	10.d	8/15/03 8:00	8/28/03 17:00
Prepare VIC 1-km datasets	10.d	8/29/03 8:00	9/11/03 17:00
Run and test 1-km datasets	30.d	9/12/03 8:00	10/23/03 17:00
Build 2.1: 1km Datasets	.d	10/23/03 17:00	10/23/03 17:00
Improve Parallel LDAS	80.d	8/1/03 8:00	11/20/03 17:00
Build 2.2: Improve Parallel LDAS	.d	11/20/03 17:00	11/20/03 17:00
Improve GDS-based I/O, Add Data Management	90.d	8/1/03 8:00	12/4/03 17:00
Build 2.3: Improve GDS I/O	.d	12/4/03 17:00	12/4/03 17:00
Implement Web-Based Interface, Activate LAS	60.d	12/5/03 8:00	2/27/04 17:00
Build 2.4: Web-based Interface	.d	2/27/04 17:00	2/27/04 17:00
Prepare Final Documentation	75.d	8/1/03 8:00	11/13/03 17:00
Prepare Final Build	5.d	12/5/03 8:00	12/11/03 17:00
Build 2.5 Final Build & Test	.d	12/11/03 17:00	12/11/03 17:00
Build 3: Second Code Improvement	.d	2/27/04 17:00	2/27/04 17:00
G) Second code improvement completed	.d	2/27/04 17:00	2/27/04 17:00
Build 4: Full interoperability demonstrated using improved codes	100.d	3/1/04 8:00	7/30/04 17:00
Integrate ESMF Utilities	20.d	3/1/04 8:00	4/1/04 17:00
Build 3.1: ESMF Utilities	.d	4/1/04 17:00	4/1/04 17:00
Improve Parallel LDAS	20.d	4/1/04 8:00	5/1/04 17:00
Build 3.2: Improve Parallel LDAS	.d	5/1/04 17:00	5/1/04 17:00
Improve GDS-based I/O	20.d	5/1/04 8:00	6/1/04 17:00
Build 3.3: Improve GDS I/O	.d	6/1/04 17:00	6/1/04 17:00
Implement Web-Based Interface	20.d	6/1/04 8:00	7/1/04 17:00
Build 3.4: Web-based Interface	.d	7/1/04 17:00	7/1/04 17:00
Prepare Final Documentation	75.d	4/9/04 8:00	7/22/04 17:00
Prepare Final Build	20.d	7/1/04 8:00	7/30/04 17:00
Build 3.5 Final Build & Test	.d	7/30/04 17:00	7/30/04 17:00
Build 4: Full interoperability demonstrated using	.d	7/30/04 17:00	7/30/04 17:00

improved codes			
J) Full interoperability demonstrated using improved codes	.d	7/30/04 17:00	7/30/04 17:00
Build 5: Customer delivery accomplished	50.d	7/01/04 17:00	8/31/04 17:00
Prepare GLDAS datasets	5.d	7/01/04 8:00	7/05/04 17:00
Run and test GLDAS datasets	5.d	7/05/04 8:00	7/10/04 17:00
Build 4.1: GLDAS build	.d	7/10/04 17:00	7/10/04 17:00
Port codes to NCEP	5.d	7/01/04 8:00	7/11/04 17:00
Prepare NCEP datasets	8.d	6/20/04 8:00	6/30/04 17:00
Run and Test NCEP build	5.d	7/12/04 8:00	7/16/04 17:00
Build 4.2: NCEP Build	.d	7/16/04 17:00	7/16/04 17:00
Prepare NLDAS datasets	5.d	7/19/04 8:00	7/23/04 17:00
Run and Test NLDAS build	5.d	7/26/04 8:00	7/30/04 17:00
Build 4.3: NLDAS Build	.d	7/30/04 17:00	7/30/04 17:00
Improve Web-Based Interface	10.d	8/2/04 8:00	8/13/04 17:00
Build 4.4: Final Web-based Interface	.d	8/13/04 17:00	8/13/04 17:00
Prepare Final Documentation	20.d	7/1/04 8:00	8/2/04 17:00
Prepare Final Build	10.d	8/16/04 8:00	8/27/04 17:00
Build 4.5 Final Build & Test	.d	8/27/04 17:00	8/27/05 17:00
Build 5: Customer delivery accomplished	.d	8/31/04 17:00	8/31/04 17:00
K) Customer delivery accomplished	.d	8/31/04 17:00	8/31/04 17:00

The initial cycle of the software development is to move the GLDAS, NOAH, and CLM to the ESS Test bed. At this stage, there will be no design changes. The expected changes will be to make the software compatible with the existing ESS Tested operating system and operating environment. This Global LDAS code baseline will be completed by July 30, 2002.

Build 1 will be the first code improvement. This is in conjunction with Milestone F that will be completed by March 30, 2003. The expected changes will be to parallelize the Global Land Data Assimilation (GLDAS) (or “driver”) portion of the code to generate better performance. This part of the code ingests data and performs spatial and temporal interpolation and gridding before initiating calls to the 1-D column land surface models. The land surface models themselves will not be parallelized. Build 1 will be on the ESS Test bed.

Build 2 will consist of generating a scheme for interoperability and community requirements and demonstrate a prototype. This is milestone I) which will be completed by July 30, 2003. Build 2 will demonstrate the VIC code on both the ESS Test bed and the LIS cluster.

Build 3 will make the second code improvement, which will add the database/data management functions and the visualization function. This will be the implementation of the GRADS/DODS servers with the Internet based user interface. The GLDAS code will

also be optimized for faster processing time, including refinements to the parallelization implemented in Build 1. All of these functions take place on the LIS cluster. This corresponds with milestone G that will be completed by Feb 30, 2003.

Build 4 was originally based on receiving an ESMF compliant version of the CLM code by June 2003. LIS now believes that a better result can be obtained by structuring the LIS driver to be a gridded component. Since LIS provides the infrastructure to drive at least three land surface models including CLM V2.0, implementing LIS itself as a gridded component in ESMF will allow CLM in LIS to be used for coupling with other Earth system models. This development will be facilitated by involvement on each other's review boards, attendance at ESMF team meetings, and at least quarterly communication. For milestone J) LIS would need to use a minimum of 3 ESMF utilities and implement the ESMF interfaces required for LIS to become a gridded component. This task will be completed by July 30, 2004.

Build 5 is the last version of the LIS under milestone K that is finished in August of 2004. For the final product, the LIS team will port the LIS code to three potential users of the LIS, NLDAS, GLDAS, and NCEP. The code will be beta tested at each platform/facility.

3.5 Code Re-use

Following development of the LIS, we plan to transition both the NLDAS and GLDAS projects to the LIS for their applications. Further, we have had discussions with NCEP about the development of an NCEP GLDAS, which would utilize the LIS. Beyond these opportunities, we seek to maximize code re-use by ensuring that our code is well documented, modular, extensible and portable.

4. Management Approach

4.1 Assumptions and Constraints

The Hydrological Sciences Branch at NASA's Goddard Space Flight Center will manage the LIS project. The software engineering and development will be managed by Dr. Christa Peters-Lidard. Each partner in the project will work from his or her home organizations' facility. Weekly status meetings will be held where any of the collaborating partners can teleconference in. Monthly status teleconferences will be held where a representative of each of the collaborating partners are encouraged to attend. Other working group meetings will be scheduled as needed.

The Land Surface Models CLM, VIC, and NOAH will not be managed by LIS. These communities have established procedures for successfully managing publicly available software. We will interface with each of these groups by having a member of the LIS team as an established member of these communities and by scheduling working group meetings as needed. Further, each of these team members will participate in monthly teleconferences to discuss land surface model issues as needed.

The GrADS/DODS team from COLA also has established procedures for managing publicly available software. COLA will be developing additional capabilities for LIS that will subsequently be available to the wider GrADS/DODS community. We will interface directly with P. Dirmeyer and B. Doty of COLA to coordinate COLA's effort, and both will participate in monthly teleconferences and additional meetings as needed.

4.2 Milestones and Schedules

These are the milestones and the not later than dates for the LIS:

Project Start	01-FEB-2002
A) Software Engineering Plan	30-JUN-2002
E) Global LDAS Code Baselined	30-JUL-2002
B) FY02 Annual Report delivered	30-AUG-2002
f) Installation of Linux Cluster	30-AUG-2002
H) Design policy for Interoperability and Community Delivery	28-FEB-2003
F) First Code Improvement	30-MAR-2003
I) Interoperability Prototype	30-JUL-2003
C) FY03 Annual Report	30-AUG-2003
G) Second Code Improvement	28-FEB-2003
J) Full Interoperability demonstrated	30-JUL-2004
K) Customer delivery Accomplished	30-AUG-2004
K) Present LIS to Review Board	30-AUG-2004
D) Final Report Delivered	28-FEB-2005

4.3 Metrics

LIS will generate metrics on the software residing on the ESS Test bed, the software on the LINUX cluster, and usage metrics on the LINUX cluster when it becomes available. Various metrics that deal with the measurement of the software product will be employed. The metrics planned for the estimating the size of the product are Total lines of code broken down into lines of executable code, comment lines, and blank lines, and number of modules (subroutines). Metrics such as cyclomatic complexity and level of nesting will be used as a measure of complexity of the code. These metrics will be reported for both the Linux cluster and the ESS Test bed. We will also report on the LIS System usage of the ESS Test bed monthly.

LIS will operate a web-based defect tracking tool. The main point of the tool will be to track any known defect to closure. The defect tracking tool will have a field to identify high, medium and low priority status for each entry. We plan to have a simple metric on number of defects per month, number of defects closed, and average time to close defects. Further information on bug tracking, fixing and validation is given in the Quality Assurance section below.

After the LIS Cluster is operational, the metrics planned are the typical metrics used for system administration and management such as system availability, processor time used,

processor time idle, number of users, etc. and other metrics to allow LIS to tune the system to be most efficient.

4.4 Risk Management

LIS plans to practice risk management to anticipate, mitigate and control risks to the project. All high and medium impact risks will be assigned to an engineer to generate a mitigation plan and track the action to closure. The mitigation plans will be approved, modified, or rejected at the next scheduled weekly status meeting and status of the mitigation plan will be an agenda item for the monthly teleconference. The risks to the LIS project will be reviewed at the beginning of each new milestone period.

A risk report will be generated and updated at the beginning of each new milestone period. The risk report will include risk name, description, level of impact, probability of impact, type of mitigation (avoid, mitigate, or accept). If mitigation action is appropriate, a brief description of the action and current status of the action is included.

The initial list of risks and risk mitigation strategies are given in Table 2.:

Table 2. Initial risks for the Land Information System

Name	Description	Level of Impact (H,M,L)	Prob. of Impact (H,M,L)	Type of Mitigation (Avoid, Mitigate, Accept)	Mitigation Actions
ESS Test bed	Availability and reliability of ESS Test bed platform	H	L	Mitigate	A monthly report on the CPU usage will be obtained to keep track of how much computer time is being used. Additionally, the baselining runs are also being done on a GSFC Code 974 Linux machines. STATUS 7/15:Special LIS queue was needed on the Origin 3000 to meet milestones.
Linux cluster	Availability and reliability of Linux cluster	H	M	Mitigate	Extensive planning and review by John Dorband's cluster team has been employed to minimize unforeseen hardware issues. In addition, extensive testing of small "prototype" clusters is ongoing and serves to identify issues in advance of the cluster due date. First year development is dependent only on the ESS Test bed in the event that the cluster is late. Status 7/15:Cluster was

					available on schedule.
ESMF	Timely availability of ESMF interface	M	L	Mitigate	LIS dependency on ESMF delivery of CLM is scheduled 3 months beyond the ESMF milestone date, to allow for late delivery. Further, reciprocal review board appointments encourage further interaction, and will be supplemented by an Interface Control Document to facilitate development and rapid implementation of the ESMF prototype CLM. Status 7/15: LIS will contact ESMF/DeLuca to get availability schedule for CLM. Status 9/15: ESMF compliant CLM not available in June 2003. LIS will propose alternate Milestone J.
Performance	Code Improvement Milestones late or not met	M-H	L	Mitigate	The LIS schedule has a goal of all milestones being completed one month before the due date. This schedule leaves time for unexpected issues. The status of the milestones are on the agenda for the monthly telecon with all of the Investigators.
Turnover	Personnel Turnover	L	M	Mitigate	A weekly status meeting will allow for each engineer to keep the group informed of their progress. Also, we have a web-based portal where we can view and archive working documents and meeting notes. Status 7/15: Web site has proven invaluable for info exchange between team members and between collaborators.

5. Product Assurance

5.1 Configuration Management

The LIS software will be managed using the tool Concurrent Versions System (CVS). CVS provides a repository for code, checkout of code, a history of changes made to the code, and a commit feature to prevent developers from making contradictory changes to the same code. The LIS test scripts and test data will also be managed using CVS.

LIS will have an assigned gatekeeper and backup gatekeeper to accept changes to the baseline version of software. All software will have version numbers. Version numbers will be based on the build number. The first release for build 1 will be Version 1.0. Any minor changes will increase the version number to Version 1.1. When using Land Surface Models such as CLM and NOAH that are in the public domain, LIS will continue using the version number that was assigned by the developing organization.

Additional documentation on CVS can be obtained from <http://www.gnu.org/manual/cvs-1.9/cvs.html>.

LIS will have a documentation repository on the public web site. The released version of each document will be displayed on the site. Non-current versions of the documentation will be stored in a simple file directory on the server. The Webmaster will control access to the web site.

5.2 Quality Assurance

The LIS will address quality in all aspects of the program. Ms. Susan Olden will be responsible for all aspects of quality assurance, and will perform the following functions:

- Ensure tests from test plans are implemented
- Assure that software defects are documented, fixed, and implemented
- Verify adherence to coding standards and code documentation requirements.
- Attend code walk-through's, and
- Ensure documentation stays current with the software and is internally consistent.

We will concentrate on conducting formal technical reviews and performing well-planned software testing. Our goal will be to prevent defects if possible, and detect and remove remaining defects as early in the life cycle of the software as possible. A web-based defect tracking tool will facilitate this rapid defect detection. The defect tool will automatically generate a message to be sent to our Configuration Management Team, at which time an entry will be logged specifying the priority of the defect. The team will then determine a suitable course of action, and depending on the severity of the defect, will report to the weekly team meeting and/or the monthly teleconference the status and outstanding issues related to the defect.

The technical review most advantageous to the LIS project will be code walkthroughs. We will walkthrough all new code generated as part of this project. The preferred method is to walkthrough small portion of the code with a small group of peers.

LIS will also conduct a Requirements review, a Software Design Review, an Interoperability Design Review, and Interface (GUI) design review, and a Test Plan Review. Each of these reviews will be held with the LIS Principal Investigators. Each review will result in review minutes that will include a list of attendees, a list of action items, and the person that is assigned to respond to the action item. The review minutes will be stored on the LIS intranet area.